

新しい社会システムデザインに向けた情報基盤技術の創出
2018 年度採択研究者

2020 年度
年次報告書

杉山 磨人

国立情報学研究所 情報学プリンシプル研究系
准教授

多変数間に潜む高次相互作用の探索と分解

§ 1. 研究成果の概要

様々な変数によって記述されるシステムの挙動を読み解くためには、変数間の相互作用の発見と、それらの分解による精緻な分析が欠かせない。例えば、身体測定データであれば、各個人は身長や体重、血圧といった変数で記述され、これらの変数間には、身長が高くなると体重が増える、のような相互作用が存在すると考えられる。しかし、多くの変数からなるより複雑な相互作用を大規模なデータから効率的に見つけ出し、その作用を適切に調べるための手法は、未だに確立されていない。本研究では、この問題を解決するための理論基盤と解析技術を構築することを目的とし、研究を進めている。

2020年度は、隠れ変数を含めた高次相互作用が及ぼす作用の分解とモデルの理論解析を主に進め、ブラインド信号源分離や点過程推定を可能とするアルゴリズムの構築に成功した。高次の相互作用を含む情報幾何学的な確率モデルを適切に構築することで、画像や音声といった信号の分解が実現できることを示した。また、量子化学分野での共同研究も進め、同様の確率モデルを利用した多体波動関数の推定について、高次相互作用を導入することで、より高精度な推定が可能となることを示した。さらに、本研究の発展として、機械学習実装の信頼性担保にも取り組んでいる。特定の機械学習アルゴリズムでは、変数間の相互作用に基づき、可能な入出力の範囲が特定されることが知られている。この性質を利用することで、この範囲から逸脱する入出力のペアが検出された際に、実装にバグが混入していると断定することができる。このアプローチの有効性を検証し、著名な機械学習手法 Lasso の実装に混入したバグを高精度で検出できることを示した。

【代表的な原著論文情報】(最大 5 件)

- 1) “Convex Optimization for Blind Source Separation on Statistical Manifolds”, Luo, S., Azizi, L., Sugiyama, M., *NeurIPS 2020 Workshop: Differential Geometry meets Deep Learning*, 1–11, 2020
- 2) “Learning Joint Intensity in a Multivariate Poisson Process on Statistical Manifolds”, Luo, S., Zhou, F., Azizi, L., Sugiyama, M., *NeurIPS 2020 Workshop: Deep Learning through Information Geometry*, 1–13, 2020.
- 3) “Artificial Neural Networks Applied as Molecular Wave Function Solvers”, Peng-Jian, Y., Sugiyama, M., Tsuda, K., Yanai, T., *Journal of Chemical Theory and Computation*, 16(6), 3513–3529, 2020
- 4) “Testing Machine Learning Code using Polyhedral Region”, Ahmed, M. S., Ishikawa, F., Sugiyama, M., *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2020)*, 1533–1536, 2020
- 5) “The Volume of Non-Restricted Boltzmann Machines and Their Double Descent Model Complexity”, Cheema, P., Sugiyama, M., *NeurIPS 2020 Workshop: Deep Learning through Information Geometry*, 1–14, 2020.